EÖTVÖS LORÁND TUDOMÁNYEGYETEM
MATEMATIKAI INTÉZET

ADAPTIVE FINITE ELEMENT METHODS FOR ELLIPTIC EQUATIONS

ADAPTÍV VÉGESELEM MÓDSZEREK ELLIPTIKUS EGYENLETEKRE

című Ph.D. értekezés tézisei

Horváth Tamás

Témavezető: Simon L. Péter
docens, PhD

Matematika Doktori Iskola
Vezető: Laczkovich Miklós
egyetemi tanár, az MTA tagja

Alkalmazott Matematika Doktori Program
Vezető: Michaletzky György
egyetemi tanár, az MTA doktora

Alkalmazott Analízis és Számításmatematikai Tanszék

Budapest, 2013

1. A dolgozat témaköre

Az elliptikus parciális differenciálegyenletek történelme több, mint 200 évre nyúlik vissza, ennek köszönhetően szerteágazó az irodalma. A vonatkozó elmélet az alapoktól a haladó szintig megtalálható pl. Evans könyvében [6]. Ugyanakkor az elliptikus parciális differenciálegyenletek általában nem oldhatók meg egzaktul, ennek megfelelően numerikusan próbáljuk közelíteni a megoldást, és igyekszünk megbecsülni a közelítés hibáját és vizsgáljuk a kvalitatív tulajdonságok megőrzését.

A különféle numerikus megoldási módszerek száma óriási, a teljesség igénye nélkül meg kell említenünk a végeselem, a véges differencia és a véges térfogat módszereket. Jelen disszertáció célja az elliptikus parciális differenciálegyenletek végeselem-módszerrel történő numerikus megoldásának vizsgálata.

A választásunk hátterében a magasfokú módszerek könnyű megkonstruálása áll. A klasszikus végeselem módszerek nagyon népszerű és jól ismert megoldási módszerek, azonban az utóbbi évtizedben a nemfolytonos végeselem-módszerek is egyre inkább elterjedtek, l. [3, 5].

Bemutatjuk ezeket a módszereket, továbbá ismertetjük az adaptív végeselem-módszereket, melyekkel a hatékonyság növelhető, és a maximum elveket, melyek fontos szerepet töltenek be, amikor elsőfokú elemeket használunk. Ezeket általában az irodalom külön kezeli, holott összekapcsolhatóak. Amikor elsőfokú elemekkel oldjuk meg a feladatot, akkor az adaptivitás során célszerű olyan rácsokat választani, melyek eleget tesznek a maximum elvekből jövő rácsfeltételeknek.

Két fő kérdéssel foglalkozunk, melyek mindegyikét további két részre bontjuk:

- Adaptivitás
  - Referencia megoldást használó módszer: egy nagyon népszerű, ugyanakkor felettébb költséges módszer. Rávilágítunk a módszer egyik hátrányára, és rámutatunk a lehetséges javításra.
  - Implicit utólagos hibabecslő: ez a módszer lokális Neumann-feladatokat old meg minden egyes résztartományon. A pontos megoldás gradiensének közelítése egy fontos részfeladat. Egy új konstrukciót fogunk mutatni, mely megadja a lehetőséget egy jobb becslés kidolgozására.
- Maximum elv
  - Összegyűjtjük a különféle maximum elv definíciókat, szükséges és elégséges feltételeket adunk azokban az esetekben, ahol ezeket még nem bizonyították korábban. Példákkal megmutatjuk, hogy bizonyos esetekben nem mindegyik őrződik meg a numerikus megoldás során.
  - Vizsgáljuk a belső büntetésen alapuló nemfolytonos végeselem-módszert egy dimenziós peremérték problémák megoldásához. Rácsfeltételeket adunk, melyek garantálják a diszkrét maximum elvet, továbbá megmutatjuk, hogy ezek a feltételek, bár csak szükségesek, valamilyen értelemben élesek is.

2. Végeselem-módszerek

A végeselem-módszert a lehető legegyszerűbb példán keresztül fogjuk bevezetni: a Laplace-egyenletet vizsgáljuk nulladrendű taggal kiegészítve, homogén Dirichlet-peremfeltétel mellett. Természetesen a peremfeltétel nem jelent valódi megszorítást, de így el tudjuk kerülni a nagyon technikai jelölésrendszert.

2.1. Gyenge megoldás

Legyen $\Omega \subset \mathbb{R}^d$ korlátos, nyílt tartomány, $\Gamma = \partial\Omega$ a pereme. Vizsgáljuk a következő másodrendű elliptikus parciális differenciálegyenletet

$$-\mathrm{div}\,(\mathcal{K}\nabla u) + \mu u = f \quad \Omega\text{−n}, \tag{1}$$

$$u = 0 \quad \Gamma\text{−n}, \tag{2}$$

ahol $\mathcal{K} : \mathbb{R}^d \to \mathbb{R}^{d\times d}$ egy szimmetrikus, egyenletesen pozitív definit mátrix értékű függvény, $\mathcal{K}_{i,j} \in L^\infty(\Omega)$ minden $i, j \in \{1, 2, \ldots, d\}$ esetén, $\mu : \mathbb{R}^d \to \mathbb{R}$ adott nemnegatív függvény, $\mu \in L^\infty(\Omega)$, $f \in L^2(\Omega)$. A klasszikus megoldás alatt egy olyan $u \in C^2(\Omega) \cap C(\overline{\Omega})$ függvényt értünk, mely pontonként eleget tesz (1)-(2)-nek.

Számos fizikai példa esetében a klasszikus megoldás létezése nem is várható el, csak az ún. gyenge megoldásé. Ezt úgy kaphatjuk meg, hogy (1)-et szorozzuk egy $v \in C^2(\Omega)$ tesztfüggvénnyel, mely a peremen eltűnik. A Green-tétel alkalmazásával a következőt kapjuk

$$a(u, v) = L(v), \tag{3}$$

ahol $a(u, v) := \int_\Omega \mathcal{K}\nabla u \cdot \nabla v + \int_\Omega \mu uv$, $L(v) := \int_\Omega fv$. A fenti egyenlet alacsonyabb regularitás mellett is értelmes, elegendő, ha $u, v \in H_0^1(\Omega)$. Az $u \in H_0^1(\Omega)$ függvényt az (1)-(2) gyenge megoldásának nevezzük, ha (3) teljesül minden $v \in H_0^1(\Omega)$ esetén.

2.2. Végeselem-módszer

Numerikusan nem tudjuk megoldani a (3) egyenletet, ugyanis $H_0^1(\Omega)$ végtelen dimenziós. Ehelyett egy jól definiált $V_{h,p} \subset H_0^1(\Omega)$ véges dimenziós altéren keressük a megoldást. Azaz keressük azt az $u_{h,p} \in V_{h,p}$ függvényt, melyre

$$a(u_{h,p}, v_{h,p}) = L(v_{h,p}), \qquad \forall v_{h,p} \in V_{h,p}. \tag{4}$$

Kihasználva $a(\cdot,\cdot)$ bilinearitását elég megkövetelni (4) teljesülését $V_{h,p}$ bázisfüggvényeire. Jelölje Φ_i ($i = 1, \ldots, N$) $V_{h,p}$ egy bázisát, ezzel $u_{h,p} = \sum_{i=1}^N c_i\Phi_i$, és a feladat a c_i együtthatók meghatározására redukálódik. Ezzel a feladatot egy lineáris egyenletrendszer megoldására vezethetjük vissza: $\mathbf{Ac} = \mathbf{L}$, ahol $(\mathbf{A})_{i,j} = a(\Phi_j, \Phi_i)$, $\mathbf{c} = (c_1, \ldots, c_N)^T$ és $(\mathbf{L})_i = L(\Phi_i)$.

Definiálnunk kell a megfelelő véges dimenziós $V_{h,p}$ teret. Mindenekelőtt felosztjuk az Ω tartományt elemekre: tipikusan háromszögekre két dimenzióban és tetraéderekre három dimenzióban. Az elemek halmazát jelölje $\mathcal{T}_h = \{E_i, i = 1, \ldots, N_{el}\}$, ahol $\cup_i E_i = \Omega$ és $int\, E_i \cap int\, E_j = \emptyset$ minden $i \neq j$ esetén. Egy további megkötésünk van a rácsra nézve: két szomszédos háromszög metszete egy teljes él kell, hogy legyen.

A bázisfüggvények definiálásának egyik módja a p-ed fokú Lagrange-bázisfüggvények használata. Ezeket megfelelően definiálva elérhetjük, hogy a polinomok folytonosak legyenek a háromszögek határain. Ezek a függvények vagy egy csúcshoz vagy egy élhez vagy egy elemhez tartoznak.

2.3. Nemfolytonos Galjorkin-módszer

Tegyük fel, hogy $u \in V_* := H_0^1 \cap H^2(\Omega)$. Ebben a szakaszban olyan véges dimenziós alteret fogunk használni, mely nem része az eredeti térnek, azaz $V_{DG} \not\subset V_*$. Ilyen esetekben a végeselem-módszert nemkonform módszernek nevezzük. Jelen esetben V_{DG} a szakaszonként p-ed fokú polinomok tere, mely polinomokról nem tesszük fel, hogy folytonosak az elemhatárokon.

Az egyszerűség kedvéért legyen $\Omega = (0,1)$ és tekintsük a következő peremérték problémát, homogén Dirichlet-peremfeltétel mellett

$$-(\kappa u')' + \mu u = f \quad \Omega\text{–n}, \tag{5}$$

ahol $\kappa, \mu \in \mathbb{R}$, $\kappa > 0$, $\mu \geq 0$.

A $\mathcal{T}_h$ rácsot a következő módon definiáljuk. Legyenek az osztópontok $0 = x_0 < x_1 < x_2 < \ldots < x_{N-1} < x_N = 1$. Vezessük be a következő jelöléseket $I_n = [x_{n-1}, x_n]$, $h_n = |I_n|$, $h_{n-1,n} = \max\{h_{n-1}, h_n\}$, (továbbá $h_{0,1} = h_1$, $h_{N,N+1} = h_N$), $\mathcal{T}_h = \{I_n, n = 1, \ldots, N\}$.

Jelöljük az egyoldali határértékeket a következő módon: $v(x_n^+) := \lim\limits_{t\to 0^+} v(x_n+t)$, $v(x_n^-) := \lim\limits_{t\to 0^+} v(x_n-t)$. Ezekkel az ugrás és az átlag egy belső pont esetén a következő képlettel definiálható

$$[\![v(x_n)]\!] := v(x_n^-) - v(x_n^+)\,, \quad \{\!\{v(x_n)\}\!\} := \frac{1}{2}(v(x_n^-) + v(x_n^+))\,.$$

A peremen az alábbi módon definiáljuk őket

$$[\![v(x_0)]\!] := -v(x_0^+)\,,\ \{\!\{v(x_0)\}\!\} := v(x_0^+)\,,\ [\![v(x_N)]\!] := v(x_N^-)\,,\ \{\!\{v(x_N)\}\!\} := v(x_N^-)\,.$$

Szorozzuk meg az (5) egyenletet egy tesztfüggvénnyel egy részintervallumon, alkalmazzuk a parciális integrálást, majd összegezzünk az egyes intervallumokra. Ezzel a következő bilineáris, illetve lineáris formákat kapjuk

$$\begin{aligned} a_{DG}(u,v) = &\sum_{n=0}^{N-1} \int_{x_n}^{x_{n+1}} \kappa u'(x) v'(x)\,\mathrm{d}x - \sum_{n=0}^{N} \{\!\{\kappa u'(x_n)\}\!\}\,[\![v(x_n)]\!] \\ &+ \varepsilon \sum_{n=0}^{N} \{\!\{\kappa v'(x_n)\}\!\}\,[\![u(x_n)]\!] + \sum_{n=0}^{N} \frac{\sigma}{h_{n,n+1}}\,[\![v(x_n)]\!]\,[\![u(x_n)]\!] + \int_0^1 \mu u(x) v(x)\,\mathrm{d}x\,, \\ L_{DG}(v) = &\int_0^1 f(x) v(x)\,\mathrm{d}x, \end{aligned} \tag{6}$$

és keressük azt az $u \in V_*$ függvényt, melyre $a_{DG}(u,v) = L_{DG}(v)$ fennáll minden $v \in V_*$ esetén. A bilineáris formában megjelenő paraméterek közül ε felel a szimmetriáért. Tetszőleges szám lehet, de általában a $\{-1, 0, 1\}$ halmazból kerül ki. A másik paraméter, σ, az ún. büntető paraméter. A bilineáris forma koercitivitása csak akkor garantálható, ha $\sigma > \sigma_0$, ahol σ_0 függ ε választásától.

Az előző szakaszban leírtakhoz hasonlóan az $a_{DG}(u,v) = L_{DG}(v)$ egyenletnek eleget tevő u függvényt numerikus nem tudjuk megtalálni a V_* térben, hiszen az végtelen dimenziós. Ehelyett definiálunk egy véges dimenziós alteret, $V_{DG} \not\subset V_*$ és abban keressük azt az u_{DG} közelítő megoldást, amire $a_{DG}(u_{DG}, v) = L_{DG}(v)$ fennáll minden $v \in V_{DG}$ esetén. A V_{DG} tér bázisának segítségével a véges dimenziós feladat ismételten visszavezethető egy lineáris egyenletrendszer megoldására.

3. Adaptivitás

A lehető legkisebb hiba elérésének egyik módja az adaptivitás használata: miután megoldjuk a feladatot egy adott rácson, utólagos hibabecslést végzünk. Ha a hiba kisebb, mint egy adott tolerancia, akkor elfogadjuk a megoldást, egyébként finomítjuk a rácsot és/vagy a használt polinomok fokait változtatjuk, és újra megoldjuk a diszkrét feladatot. Az utólagos (vagy a-poszteriori) hibabecslés annyiban különbözik az a-priori hibabecsléstől, hogy csak a rendelkezésre álló adatokat használja, azaz nem tartalmazza az ismeretlen u megoldást.

Négy fő változata létezik az adaptivitásnak: h-adaptivitás (a rács változik), p-adaptivitás (a polinomfok változik), hp-adaptivitás (a rács és a polinomfok változik), r-adaptivitás (a rácspontok helyzete változik).

Az adaptív végeselem-módszerek az alábbi séma alapján működnek:

Kezdet: megoldjuk a problémát alacsony polinomfok mellett egy ritka rácson

Ismételjük az alábbit:

S1 megbecsüljük a hibát,

S2 ha a hiba elég kicsi, akkor megállunk,

S3 egyébként megállapítjuk, hogy hol és hogyan finomítjuk/durvítjuk a rácsot és a polinomfokot,

S4 kiszámoljuk az új megoldást és S1-re ugrunk,

A módszerek közti fő különbség a hibabecslésben és a finomításban rejlik. A durvítás azt jelenti, hogy a polinomfokot csökkentjük vagy pár rácselemet összeolvasztunk. Azokon az elemeken alkalmazzuk ezt, ahol a hiba kicsi. A durvítás célja az ismeretlenek számának kontrollálása.

3.1. Referencia megoldást használó hp-adatív módszer

A referencia megoldást használó módszer részletes leírása a [7] és [8] cikkekben található. Hibabecslésként az alábbi alapötletet használja az S1 és S2 lépésekhez:

S1a kiszámoljuk $u_{h,p} \in V_{h,p}$-t (4) megoldásával,

S1b kiszámoljuk a referencia megoldást $u_{h/2,p+1} \in V_{h/2,p+1}$ megoldva (4)-et a bővebb $V_{h/2,p+1}$ térben,

S1c használjuk $u_{h/2,p+1}$-et mint egy pontosabb megoldást. Ennek segítségével definiáljuk a $u_{h,p} - u_{h/2,p+1}$ hiba indikátort,

$$1)\ \eta_T = \frac{|u_{h,p} - u_{h/2,p+1}|_{H^1(T)}}{|u_{h/2,p+1}|_{H^1(T)}} \qquad 2)\ \eta_T = \frac{\|u_{h,p} - u_{h/2,p+1}\|_{H^1(T)}}{\|u_{h/2,p+1}\|_{H^1(T)}}$$

$$3)\ \eta_T = |u_{h,p} - u_{h/2,p+1}|_{H^1(T)} \qquad 4)\ \eta_T = \|u_{h,p} - u_{h/2,p+1}\|_{H^1(T)}$$

S2 ha $\sqrt{\sum_T \eta_T^2} < \mathtt{TOL}$, akkor megállunk.

Ahol az alábbi jelöléseket használtuk: $|u|^2_{H^1(T)} := \int_T \sum_{i=1}^d |\partial_i u|^2$ és $\|u\|^2_{H^1(T)} := |u|^2_{H^1(T)} + \int_T |u|^2$.

A különbség [7] és [8] között η_T kiválasztásában rejlik, illetve abban, hogy [7] az élek mentén is elvégez egyfajta finomítást.

Kövessük a [9]-ben leírt ötletet és tegyük fel, hogy találunk egy olyan $f \neq 0$ függvényt, melyre $\forall v_{h,p} \in V_{h,p}$ esetén $\int_\Omega f v_{h,p} = 0$. Ilyen esetben (4) a következőre redukálódik

$$a(u_{h,p}, v_{h,p}) = 0,$$

aminek a megoldása $u_{h,p} = 0$. Ugyanakkor, ha $f \neq 0$, akkor a pontos megoldás is különbözik nullától.

Ha $\forall v_{h/2,p+1} \in V_{h/2,p+1}$ esetén $\int_\Omega f v_{h/2,p+1} = 0$, akkor kihasználva a $V_{h,p} \subset V_{h/2,p+1}$ összefüggést, kapjuk, hogy $\int_\Omega f v_{h,p} = 0$ is fennáll. Ebben az esetben $u_{h,p} = u_{h/2,p+1} = 0$ és így a becsült hiba nulla, annak ellenére, hogy a valódi hiba tetszőleges lehet.

Megmutatjuk, hogy találhatunk egy, a célnak megfelelő f függvényt. Adott $T \in \mathcal{T}_h$ esetén definiáljuk az $u_c : \Omega \to \mathbb{R}$ függvényt úgy, hogy $\mathrm{supp}(u_c) = T$, $u_c(x,y) = \chi_T(x,y) b_T^2(x,y) p(x,y)$, ahol χ_T a T karakterisztikus függvénye, $b_T : \Omega \to \mathbb{R}$ olyan T-n értelmezett függvény, melyre $b_T = 0$ a T peremén, a p polinomot pedig a következő alakban keressük

$$p(x,y) = \sum_{k=0}^{m-1} c_k x^{a_k} y^{b_k},$$

ahol $a_k, b_k \in \mathbb{N} \cup \{0\}$ tetszőleges kitevők. Fel tudunk állítani egy lineáris egyenletrendszert, aminek megoldásaként megkapjuk az együtthatókat. William F. Mitchellnek köszönhetően tesztelhettük a módszert a PHAML programcsomaggal.

Könnyen ki tudjuk javítani a problémát egy tartalék hibabecslő beépítésével. Például használhatjuk a reziduális-alapú hibabecslőt

$$|||u - u_{h,p}|||^2 \leq \eta_{\mathrm{res}}^2 = C_{\mathrm{res}} \left(\sum_{T \in \mathcal{T}_h} h_T^2 \|r\|_{L_2(T)}^2 + \sum_{\gamma \in \partial T} h_T \|R\|_{L_2(\gamma)}^2 \right), \tag{7}$$

ahol $r = f + \mathrm{div}\,(K \nabla u_{h,p}) - \mu u_{h,p}$, $R = \left[\frac{\partial u_{h,p}}{\partial \eta}\right]$ a numerikus megoldás ugrása a belső élen, $|||u|||^2 = a(u,u)$ pedig az energia-norma (l. [2] a részletekért), a konstans C_{res} független h-tól.

Közismert, hogy a referencia megoldást használó módszer finomítási technikája kiváló, l. [8], ugyanakkor (7) egy garantált felső becslést ad a hibára, de ennek a finomítási technikája sokkal bonyolultabb. Ezeket figyelembe véve a következőképpen módosítjuk az S2 lépést:

S2 ha a hiba kicsi, alkalmazzuk (7)-t

- S2a ha $\eta_{\mathrm{res}} < \mathtt{TOL}$, leáll az algoritmus,
- S2b egyébként egy teljes finomítást végzünk (finomítjuk a rácsot és emeljük a polinomfokot) és az S4 lépéssel folytatjuk.

3.2. Implicit hibabecslés magasfokú illesztéssel

Az (1)-(2) egyenletet fogjuk vizsgálni $\mathcal{K} = I$, $0 < \mu \in \mathbb{R}$ mellett. $\Omega \subset \mathbb{R}^d$, $f \in L^2(\Omega)$, $\Gamma = \partial\Omega$ az Ω pereme. Ezekkel a jelölésekkel a feladat

$$-\Delta u + \mu u = f \quad \Omega\text{−n}, \tag{8}$$

$$u = 0 \quad \Gamma\text{−n}. \tag{9}$$

Az implicit hibabecslő az $e_{h,p} := u - u_{h,p}$ hibára minden $T \in \mathcal{T}_h$ résztartományon felírható egyenleten alapszik. Felhasználva, hogy $-\Delta e_{h,p} + \mu e_{h,p} = -\Delta(u - u_{h,p}) + \mu(u - u_{h,p})$, a következőt kapjuk

$$-\Delta e_{h,p} + \mu e_{h,p} = f - (-\Delta u_{h,p} + \mu u_{h,p}) \qquad T\text{−n}, \tag{10}$$

$$\partial_\nu e_{h,p} = \partial_\nu (u - u_{h,p}) = \partial_\nu u - \partial_\nu u_{h,p} \quad \partial T\text{−n}. \tag{11}$$

A peremfeltétel jobb oldala ismeretlen. Ennek megfelelően valahogy közelítenünk kell. Az első ötlet a $\partial_\nu u_{h,p}$ értékek átlagolása: az élet tartalmazó két háromszög mindegyikén kiszámoljuk $\partial_\nu u_{h,p}$-t, és ezt átlagoljuk.

Ha p-ed fokú polinomokkal oldjuk meg a feladatot, akkor az átlagolás eredményeképp a $\partial_\nu u$-ra kapott közelítés $p-1$-ed fokú polinom. Ugyanakkor pl. [1]-ben is azt találjuk, hogy a lokális feladatot magasabb fokú polinomokkal kell megoldani, mint az eredeti feladatot. Ennek megfelelően egy magasabb fokú közelítés kell a Neumann-peremfeltételhez is.

Olyan $\hat{e}_{h,p}$ hibabecslőt fogunk konstruálni, mely $p+1$-ed fokú minden $T \in \mathcal{T}_h$ résztartomány esetén. Egy megfelelő

$$G_{p,T}(u_{h,p}) \approx \nabla u_T$$

közelítés esetén az $\hat{e}_{h,p}$ hibabecslő a (10)-(11) alapján az alábbi egyenlet végeselem megoldása

$$-\Delta \hat{e}_{h,p} + \mu \hat{e}_{h,p} = f + \Delta u_{h,p} - \mu u_{h,p} \qquad T\text{−n}, \tag{12}$$

$$\partial_\nu \hat{e}_{h,p} = \nu \cdot G_{p,T}(u_{h,p}) - \partial_\nu u_{h,p} \quad \partial T\text{−n}. \tag{13}$$

A következő extra feltétel garantálja a szuperkonvergenciát. Tegyük fel, hogy létezik egy olyan $C(u)$ konstans, amihez van olyan $\tau \geq 0$, melyre

$$\|\nabla(u_{h,p} - \mathcal{I}_{h,p} u)\|_0 \leq C(u) h^{p+\tau} \tag{SC}$$

tetszőleges $h > 0$ esetén (h a rácselemek átmérői közül a legnagyobb).

Tekintsük az alábbi diszkrét gradiens operátort

$$G_{p,T} : W^{1,\infty}(\tilde{T}) \to [L^1(T)]^d,$$

ahol p jelzi a lokális polinomfok függését a $V_{h,p}$ végeselem tértől, $\tilde{T} = T \cup T_1 \cup \ldots \cup T_k$, ahol T és T_i szomszédosak ($i = 1, \ldots, k$). Ennek megfelelően definiáljuk a $G_p : W^{1,\infty}(\mathcal{T}_h) \to [L^1(\mathcal{T}_h)]^d$ operátort úgy, hogy $G_p|_T = G_{p,T}$ minden $T \in \mathcal{T}_h$ esetén. A következő négy feltételből az első hármat a [2] könyvből vettük át, míg a negyedig saját feltétel, és sok helyen ez határozza meg az analízist:

(A1) $G_{p,T}(v)$ csak $v|_{\tilde{T}}$-tól függ,

(A2) $G_{p,T} : W^{1,\infty}(\tilde{T}) \to [L^1(T)]^d$ folytonos,

(A3) Ha $u \in \mathcal{P}_{p+1}(\tilde{T})$, akkor $G_{p,T}(\mathcal{I}_{h,p,\tilde{T}} u) = \mathcal{I}_{h,p,T} \nabla u_T$,

(A4) $G_{p,T}(u_{h,p})$ egy gradiens, azaz $\exists \mathcal{G}_p(u_{h,p}) \in W^{1,\infty}(\Omega)$, melyre $G_{p,T}(u_{h,p}) = \nabla \mathcal{G}_p(u_{h,p})|_T$.

1. Tétel. *(H., Izsák [10]) Tegyük fel, hogy $(A1), (A2), (A3), (A4)$ és (SC) fennáll. Ekkor létezik olyan $C_1 > 0$, melyre a* (12)-(13) *által meghatározott hibabecslő eleget tesz a következőnek*

$$\sum_{T \in \mathcal{T}_h} \|e_{h,p} - \hat{e}_{h,p}\|_{1,T}^2 \leq C_1 C^2(u) h^{2(p+\tau)} |u|_{p+2}^2,$$

ahol C_1 függ μ-től és C_P-től, ami a Poincare-egyenlőtlenségből származik.

4. Folytonos és diszkrét maximum elvek elliptikus operátorok esetén

Előrebocsátjuk, hogy a maximum elveket operátorokra fogalmazzuk meg egyenletek helyett. Legyen $\Omega \subset \mathbb{R}^d$ korlátos, nyílt tartomány, $\partial\Omega$ a pereme, $\overline{\Omega} = \Omega \cup \partial\Omega$ pedig a lezártja. A következő $\mathcal{A}$ elliptikus operátort vizsgáljuk, $\operatorname{dom}\mathcal{A} = C^2(\Omega) \cap C(\overline{\Omega})$

$$\mathcal{A}u = -\sum_{i,j=1}^{d} \frac{\partial}{\partial x_j}\left(\mathcal{K}_{ij}\frac{\partial u}{\partial x_i}\right) + \mu u, \tag{14}$$

ahol $\mathcal{K}_{ij} \in C^1(\Omega), 0 \le \mu \in C(\Omega)$. Jegyezzük meg, hogy az együttható függvények simasága lehetővé tenné az operátor átírását divergenciamentes alakra, de a fenti megközelítés kényelmesebbé teszi a maximum elves vizsgálódásokat.

2. Definíció. *Azt mondjuk, hogy az $\mathcal{A}$ operátor rendelkezik*

- *a gyenge maximum elvvel (wMP), ha fennáll a következő $\mathcal{A}u \le 0$ Ω-n $\Rightarrow \max_{\overline{\Omega}} u \le \max\{0, \max_{\partial\Omega} u\}$;*
- *az erős maximum elvvel (sMP), ha rendelkezik a gyenge maximum elvvel, továbbá ($\mathcal{A}u \le 0$ Ω-n és $\max_\Omega u = \max_{\overline{\Omega}} u = m \ge 0) \Rightarrow u \equiv m$ $\overline{\Omega}$-n;*
- *a szigorú gyenge maximum elvvel (WMP), ha $\mathcal{A}u \le 0$ Ω-n $\Rightarrow \max_{\partial\Omega} u = \max_{\overline{\Omega}} u$;*
- *a szigorú erős maximum elvvel (SMP), ha rendelkezik a szigorú gyenge maximum elvvel, továbbá ($\mathcal{A}u \le 0$ Ω-n és $\max_\Omega u = \max_{\overline{\Omega}} u = m) \Rightarrow u \equiv m$ $\overline{\Omega}$-n.*

A (14) operátorra ismert az alábbi tétel.

3. Tétel. *Ha a (14) képlet által definiált $\mathcal{A}$ operátor egyenletesen elliptikus (létezik olyan $\lambda > 0$, melyre $\lambda\|\xi\|_E^2 \le \mathcal{K}(x)\xi \cdot \xi$ minden $x \in \Omega$, $\xi \in \mathbb{R}^d$ esetén, ahol $\|\cdot\|_E$ az euklideszi norma az $\mathbb{R}^d$ téren) és*

- *$\mu \ge 0$, akkor rendelkezik a wMP-vel;*
- *$\mu \ge 0$, továbbá, Ω összefüggő, akkor rendelkezik a sMP-vel;*
- *$\mu = 0$, akkor rendelkezik a WMP-vel;*
- *$\mu = 0$, továbbá, Ω összefüggő, akkor rendelkezik a SMP-vel.*

A következő lépés a rács definiálása Ω-n. Az egydimenziós rács intervallumokból áll. A diszkrét maximum elv esetén az irodalom két dimenzióban a háromszögrácsokra vagy a hibrid rácsokra (háromszögek és négyszögek) koncentrál, míg három dimenzió esetén a tetraéder vagy téglatest rácsok az érdekesek. A rács meghatározza a $\mathcal{X} = \{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_N\}$ és $\mathcal{X}_\partial = \{\mathbf{x}_{N+1}, \mathbf{x}_{N+2}, \ldots, \mathbf{x}_{N+N_\partial}\}$ halmazokat, melyek az Ω-n illetve a $\partial\Omega$-n lévő rácspontok. Vezessük be a következő jelöléseket: $\overline{N} = N + N_\partial$ és $\overline{\mathcal{X}} = \mathcal{X} \cup \mathcal{X}_\partial$.

Ezek után készen állunk a diszkrét maximum elvek definiálására. A $V_{h,1}$ teret fogjuk használni, ugyanis ennek létezik olyan bázisa, mely rendelkezik a következő tulajdonságokkal:

- $\Phi_i(\mathbf{x}) \ge 0$ fennáll minden $\mathbf{x} \in \overline{\Omega}$ és $i = 1, \ldots \overline{N}$ esetén;
- a bázisfüggvények lineáris kombinációjában az együtthatók reprezentálják a lineáris kombináció eredményeképp kapott függvény értékét az $\overline{\mathcal{X}}$ pontjaiban.

Végül el tudjuk készíteni az ún. merevségi mátrixot: $\mathbf{A} \in \mathbb{R}^{N \times \overline{N}}$ $\mathbf{A}_{ij} = a(\Phi_j, \Phi_i)$ ez a mátrix lesz a (14) operátor diszkrét megfelelője.

A későbbiekre való tekintettel vezessük be a következő felosztott formát: $\mathbf{A} = [\mathbf{A}_0 | \mathbf{A}_\partial]$, ahol $\mathbf{A}_0 \in \mathbb{R}^{N \times N}$, $\mathbf{A}_\partial \in \mathbb{R}^{N \times N_\partial}$, amit egy $\mathbf{u} = [\mathbf{u}_0 | \mathbf{u}_\partial]^T \in \mathbb{R}^{\overline{N}}$ vektorra alkalmazhatunk, ahol $\mathbf{u}_0 \in \mathbb{R}^N$, $\mathbf{u}_\partial \in \mathbb{R}^{N_\partial}$. A konstrukció lényege, hogy külön kezeljük a belső pontokat és a perempontokat. Feltesszük, hogy $N, N_\partial \geq 2$.

4. Definíció. *Azt mondjuk, hogy az $\mathbf{A}$ mátrix rendelkezik*

- *a diszkrét gyenge maximum elvvel (DwMP), ha* $\mathbf{Au} \leq \mathbf{0} \Rightarrow \max \mathbf{u} \leq \max\{0, \mathbf{u}_\partial\}$;
- *a diszkrét erős maximum elvvel (DsMP), ha rendelkezik a diszkrét gyenge maximum elvvel, továbbá fennáll, hogy (*$\mathbf{Au} \leq \mathbf{0}$ *és* $\max \mathbf{u} = \max \mathbf{u}_0 = m \geq 0) \Rightarrow \mathbf{u} = m\mathbf{e}$;
- *a diszkrét szigorú gyenge maximum elvvel (DWMP), ha* $\mathbf{Au} \leq \mathbf{0} \Rightarrow \max \mathbf{u}_\partial = \max \mathbf{u}$;
- *a diszkrét szigorú maximum elvvel (DSMP), ha rendelkezik a diszkrét szigorú gyenge maximum elvvel, továbbá fennáll, hogy (*$\mathbf{Au} \leq \mathbf{0}$ *és* $\max \mathbf{u} = \max \mathbf{u}_0 = m) \Rightarrow \mathbf{u} = m\mathbf{e}$;

ahol $\mathbf{e}$ azt a vektort jelöli, melynek minden koordinátája 1.

A következőben szükséges és elégséges feltételeket adunk a fentiekre. A Tétel második és negyedik részét bizonyítottuk [12]-ben.

5. Tétel. *(Mincsovics, H. [12]) Az $\mathbf{A}$ mátrix rendelkezik*

- *a DwMP-vel pontosan akkor, ha a következő három feltétel teljesül*

  (w1) $\mathbf{A}_0^{-1} \geq \mathbf{0}\,;$ (w2) $-\mathbf{A}_0^{-1}\mathbf{A}_\partial \geq \mathbf{0}\,;$ (w3) $-\mathbf{A}_0^{-1}\mathbf{A}_\partial\mathbf{e} \leq \mathbf{e}\,.$

- *a DsMP-vel pontosan akkor, ha a következő három feltétel teljesül*

  (s1) $\mathbf{A}_0^{-1} > \mathbf{0}\,;$ (s2) $-\mathbf{A}_0^{-1}\mathbf{A}_\partial > \mathbf{0}\,;$ (s3) $-\mathbf{A}_0^{-1}\mathbf{A}_\partial\mathbf{e} < \mathbf{e}$ *vagy* $-\mathbf{A}_0^{-1}\mathbf{A}_\partial\mathbf{e} = \mathbf{e}\,.$

- *a DWMP-vel pontosan akkor, ha a következő három feltétel teljesül*

  (W1) $\mathbf{A}_0^{-1} \geq \mathbf{0}\,;$ (W2) $-\mathbf{A}_0^{-1}\mathbf{A}_\partial \geq \mathbf{0}\,;$ (W3) $-\mathbf{A}_0^{-1}\mathbf{A}_\partial\mathbf{e} = \mathbf{e}\,.$

- *a DSMP-vel pontosan akkor, ha a következő három feltétel teljesül*

  (S1) $\mathbf{A}_0^{-1} > \mathbf{0}\,;$ (S2) $-\mathbf{A}_0^{-1}\mathbf{A}_\partial > \mathbf{0}\,;$ (S3) $-\mathbf{A}_0^{-1}\mathbf{A}_\partial\mathbf{e} = \mathbf{e}\,.$

A fenti tétel egy elméleti eredmény, ebben a formában nehéz alkalmazni. Ehelyett például a (w1)-(w3) feltételek a következő, gyakorlatban jobban használható, ám csak szükséges feltételekre cserélhetők.

6. Tétel. *([4]) Az $\mathbf{A}$ mátrix rendelkezik a diszkrét gyenge maximum elvvel, ha a következő feltételek teljesülnek:*

(P1) $\mathbf{A}_0$ *egy nemszinguláris M-mátrix;* (P2) $-\mathbf{A}_\partial \geq \mathbf{0}$; (P3) $\mathbf{Ae} \geq \mathbf{0}$.

4.1. Diszkrét gyenge maximum elv IPDG operátorokra

Legyen $\Omega = (0, 1)$ és tekintsük a következő $\mathcal{A}$ elliptikus operátort, $\operatorname{dom} \mathcal{A} = C^2(\Omega) \cap C(\overline{\Omega})$

$$\mathcal{A}u = -(\kappa u')' + \mu u\,, \tag{15}$$

ahol $\kappa, \mu \in \mathbb{R}$, $\kappa > 0$, $\mu \geq 0$. Ez az operátor rendelkezik a gyenge maximum elvvel a 3. Tétel alapján.

Diszkretizáljuk a feladatot az IPDG módszerrel. A bilineáris formát, a_{DG}-t a (6) képlet definiálja. Vezessük be a a_{DG}^{hD} formát, mely külsőleg ugyanaz, mint (6), de csak azokon a polinomokon értelmezzük, melyek nullát vesznek fel $x = 0$ és $x = 1$ esetén.

7. Tétel. *(H., Mincsovics [11]) Legyen* $\mathbf{A} = [\mathbf{A_0}|\mathbf{A}_\partial]$ *a merevségi mátrix, amit a* (15) *operátor* a_{DG}*-vel történő diszkretizálásával kaptunk. Ez a mátrix akkor rendelkezik a diszkrét gyenge maximum elvvel, ha a paramétereket a következőképpen választjuk meg:*

- ε*-t úgy, hogy* $-\dfrac{1}{2} \leq \varepsilon \leq 0$*, ha* $\mu = 0$ *vagy* $-\dfrac{1}{2} < \varepsilon \leq 0$*, ha* $\mu > 0$*,*
- σ*-t úgy, hogy* $\dfrac{\kappa(1-\varepsilon)}{2} \leq \sigma$*,*
- *a* $\mathcal{T}_h$ *rácsot úgy, hogy* $h_i^2 \leq \dfrac{3\kappa(2\varepsilon+1)}{\mu}$*,* $i = 1, N$ *(finomság a peremen) és* $h_i^2 \leq \dfrac{3\kappa(\varepsilon+1)}{\mu}$*,* $i = 2, \ldots, N-1$*, (finomság a belső intervallumok esetén). Végül* $\mathcal{T}_h$ *tegyen eleget annak, hogy*
$$\frac{h_{i,i+1}}{h_{i+1}} - \frac{\varepsilon h_{i,i+1}}{h_i} \leq \frac{2\sigma}{\kappa}, \quad \text{illetve} \quad \frac{h_{i,i+1}}{h_i} - \frac{\varepsilon h_{i,i+1}}{h_{i+1}} \leq \frac{2\sigma}{\kappa}\,, \quad i = 1, \ldots, N-1\,. \quad \text{(egyenletesség)}$$

8. Tétel. *(H., Mincsovics [11]) Legyen* $\mathbf{A} = \mathbf{A_0}$ *a merevségi mátrix, amit a* (15) *operátor* a_{DG}^{hD}*-vel történő diszkretizálásával kaptunk. Ez a mátrix akkor rendelkezik a diszkrét gyenge maximum elvvel, ha a paramétereket a következőképpen választjuk meg:*

- ε*-t úgy, hogy* $-1 \leq \varepsilon \leq 0$*, ha* $\mu = 0$ *vagy* $-1 < \varepsilon \leq 0$*, ha* $\mu > 0$*,*
- σ*-t úgy, hogy* $\dfrac{\kappa(1-\varepsilon)}{2} \leq \sigma$*,*
- *a* $\mathcal{T}_h$ *rácsot úgy, hogy* $h_i^2 \leq \dfrac{3\kappa(\varepsilon+1)}{\mu}$*,* $i = 2, \ldots, N-1$*, (finomság a belső intervallumok esetén). Végül* $\mathcal{T}_h$ *tegyen eleget annak, hogy*
$$\frac{h_{i,i+1}}{h_{i+1}} - \frac{\varepsilon h_{i,i+1}}{h_i} \leq \frac{2\sigma}{\kappa}, \quad \text{illetve} \quad \frac{h_{i,i+1}}{h_i} - \frac{\varepsilon h_{i,i+1}}{h_{i+1}} \leq \frac{2\sigma}{\kappa}\,, \quad i = 1, \ldots, N-1\,. \quad \text{(egyenletesség)}$$

A [11] cikkben megmutattuk, hogy ezek a feltételek valamilyen értelemben élesek.

Hivatkozások

[1] M. Ainsworth. The influence and selection of subspaces for a posteriori error estimators. *Numer. Math.*, 73(4):399–418, 1996.

[2] M. Ainsworth and J. T. Oden. *A posteriori error estimation in finite element analysis*. Pure and Applied Mathematics (New York). Wiley-Interscience [John Wiley & Sons], New York, 2000.

[3] D. N. Arnold, F. Brezzi, B. Cockburn, and L. D. Marini. Unified analysis of discontinuous Galerkin methods for elliptic problems. *SIAM J. Numer. Anal.*, 39(5):1749–1779, 2001/02.

[4] P. G. Ciarlet. Discrete maximum principle for finite-difference operators. *Aequationes Math.*, 4:338–352, 1970.

[5] D. A. Di Pietro and A. Ern. *Mathematical aspects of discontinuous Galerkin methods*, volume 69 of *Mathématiques & Applications (Berlin) [Mathematics & Applications]*. Springer, Heidelberg, 2012.

[6] L. C. Evans. *Partial differential equations*, volume 19 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, second edition, 2010.

[7] L. Demkowicz, W. Rachowicz, and P. Devloo. A fully automatic hp-adaptivity. In *Proceedings of the Fifth International Conference on Spectral and High Order Methods (ICOSAHOM-01) (Uppsala)*, volume 17, pages 117–142, 2002.

[8] P. Šolín, J. Červený, and I. Doležel. Arbitrary-level hanging nodes and automatic adaptivity in the hp-fem. *Math. Comput. Simulation*, 77:117–132, 2008.

A jelölt publikációi

[9] T. L. Horváth. A note on reference solution based hp-adaptive methods. unpublished.

[10] T. L. Horváth and F. Izsák. Implicit a posteriori error estimation using patch recovery techniques. *Cent. Eur. J. Math.*, 10(1):55–72, 2012.

[11] T. L. Horváth and M. E. Mincsovics. Discrete maximum principle for interior penalty discontinuous Galerkin methods. *Cent. Eur. J. Math.*, 11(4):664–679, 2013.

[12] M. E. Mincsovics and T. L. Horváth. On the differences of the discrete weak and strong maximum principles for elliptic operators. In *Large-scale scientific computing*, volume 7116 of *Lecture Notes in Comput. Sci.*, pages 614–621. Springer, Heidelberg, 2012.

A jelölt további publikációi

[13] Á. Csík, T. L. Horváth, P. Földesi An approximate Analytic Solution of the Inventory Balance Delay Differential Equation. *Acta Technica Jaurinensis Series Logistica*, 3:231–256, 2010.

[14] T. L. Horváth and P. L. Simon. On the exact number of solutions of a singular boundary-value problem. *Differential and Integral Equations*, 22(7-8):787–796, 2009.